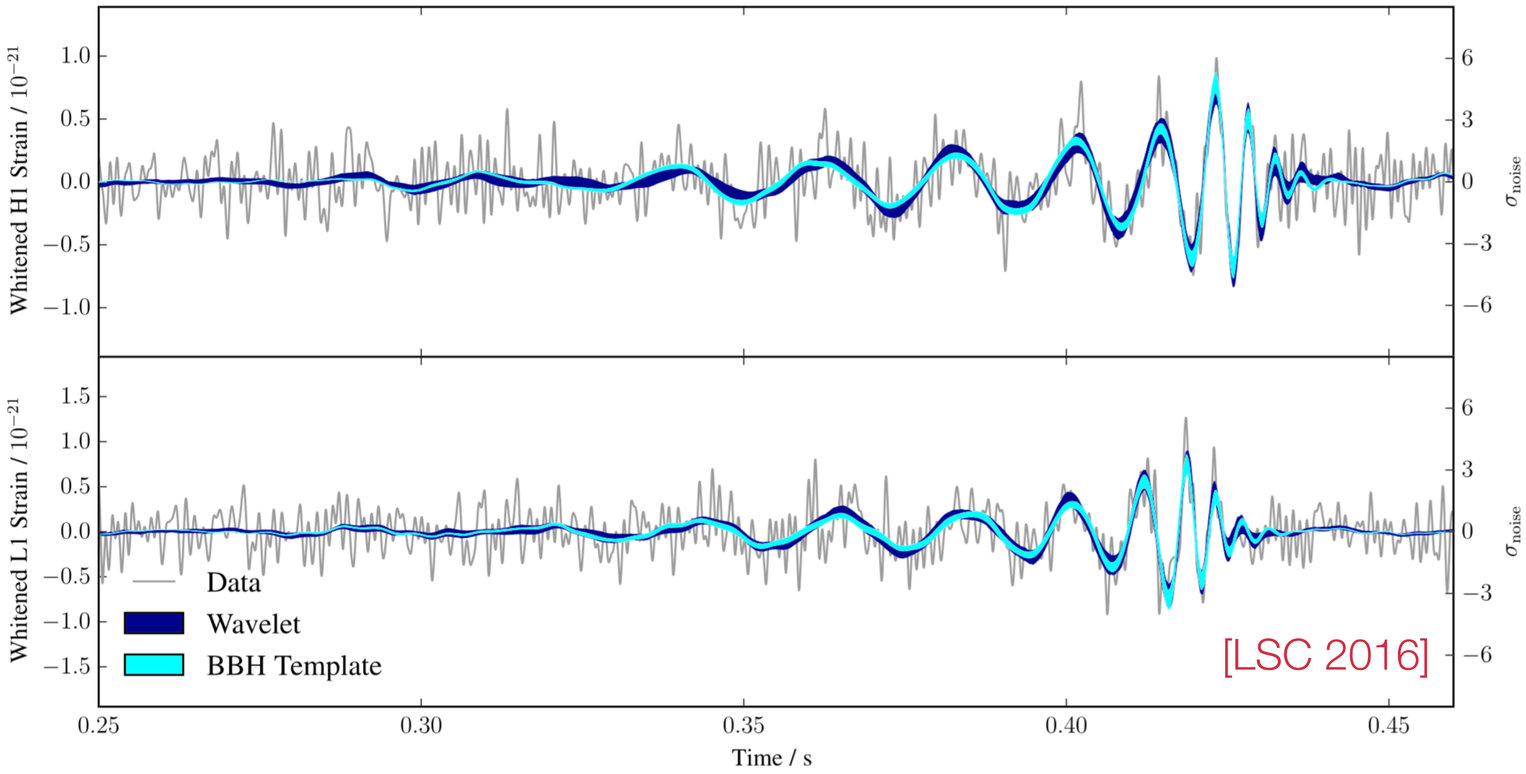
Testing general relativity with black-hole binary observations: a manifesto

Michele Vallisneri
Jet Propulsion Laboratory, Caltech

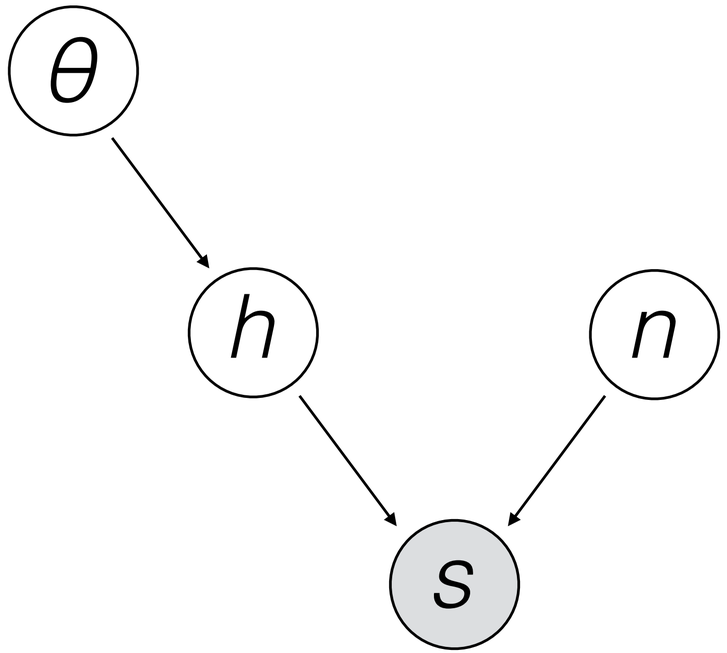


- Detecting GWs offered a profound confirmation of general relativity
- Using GW observations to test GR to increasing precision, we hope to obtain clues of new physics, and ultimately prove a more general theory

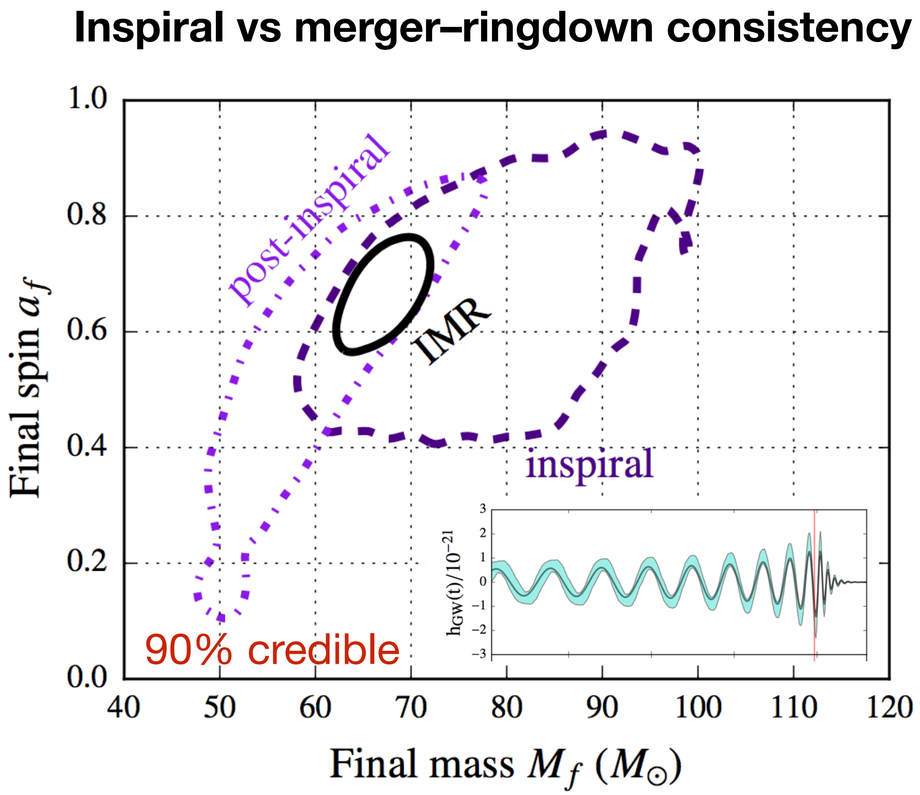
Testing GR with GW observations: a manifesto

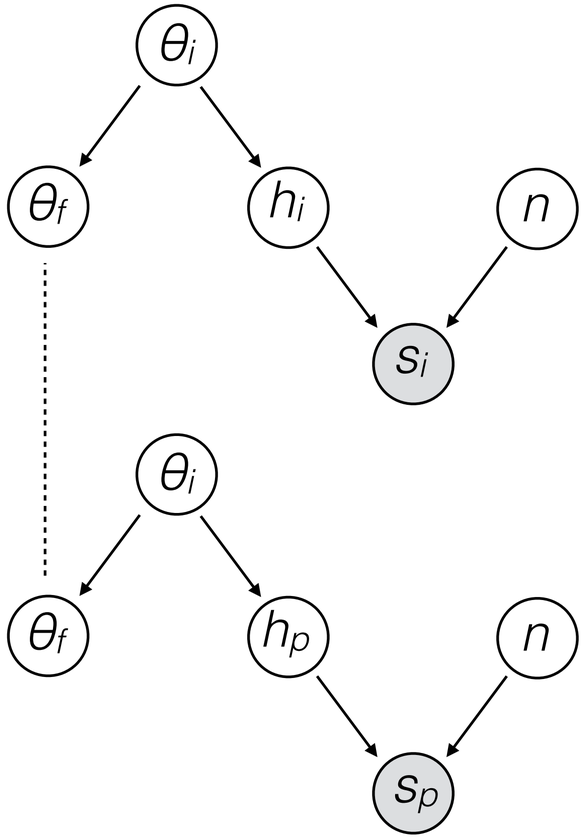
- **Consistency**: a useful sanity check, but hard to interpret statistically. P values are possible with much work. But would we ever believe an inconsistent result?

the graphical model of parameter inference

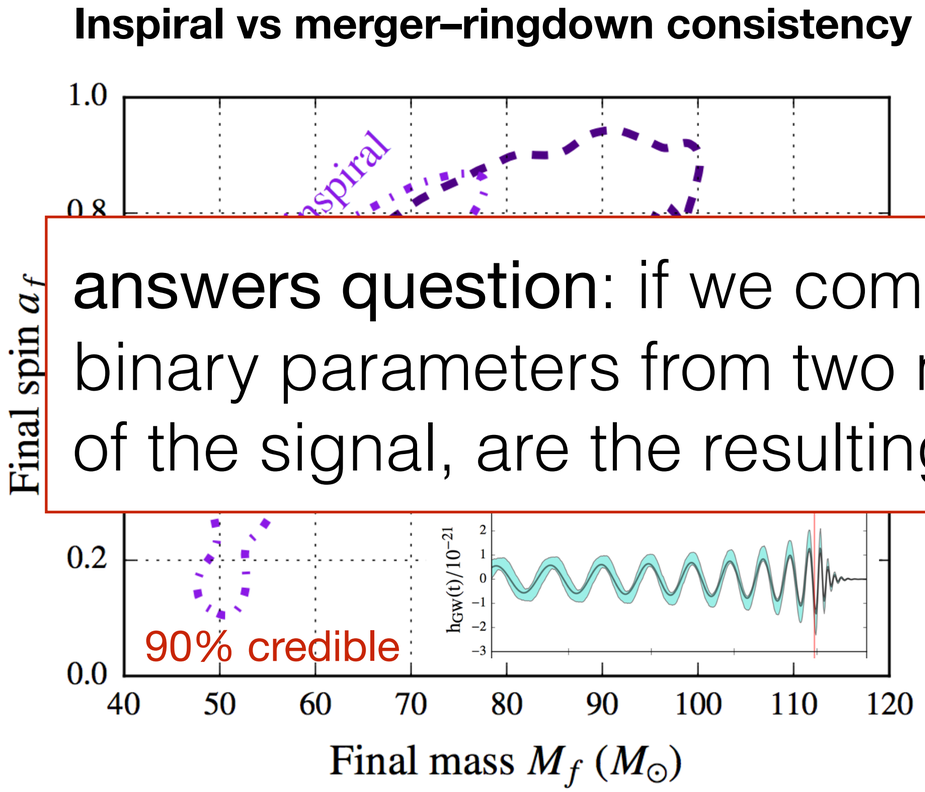


GW150914 tests of “consistency”: partial waveforms (1)

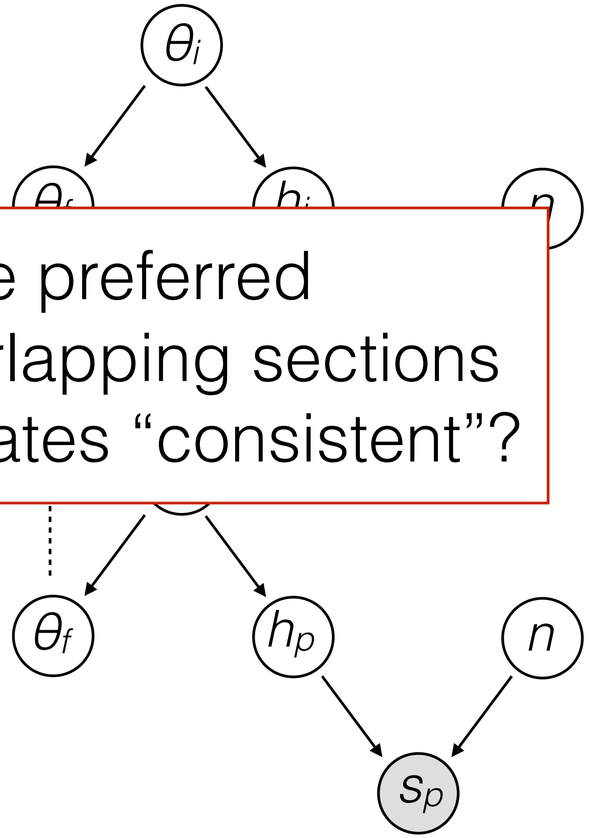




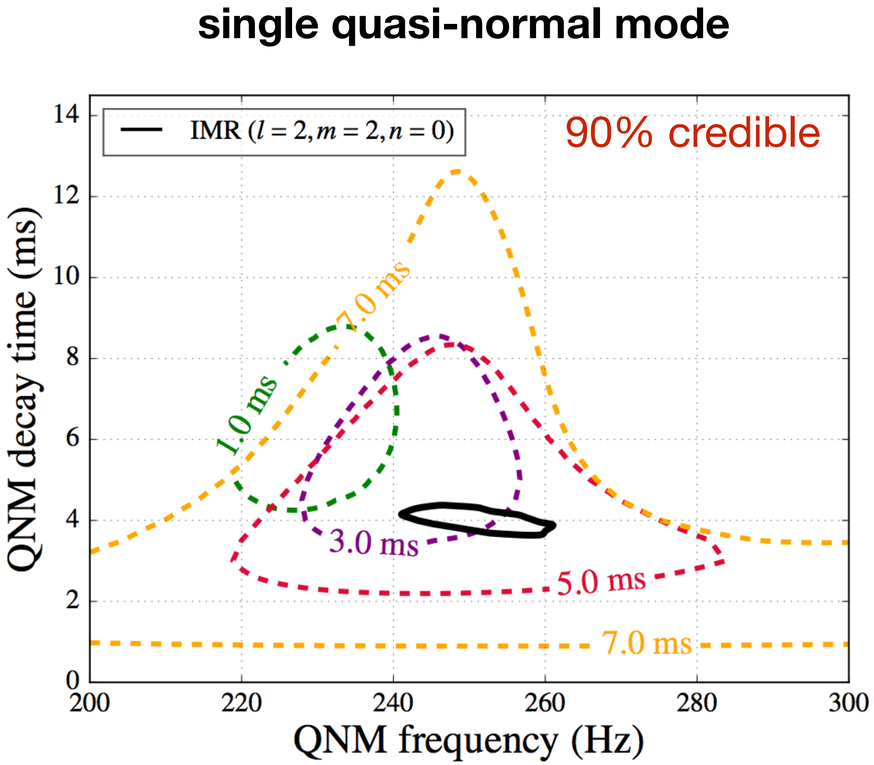
GW150914 tests of “consistency”: partial waveforms (1)

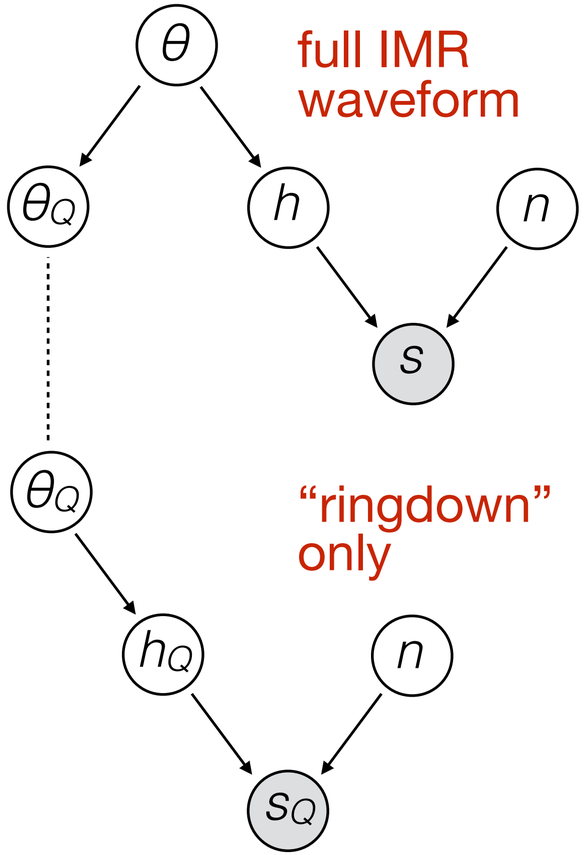


answers question: if we compute the preferred binary parameters from two nonoverlapping sections of the signal, are the resulting estimates “consistent”?



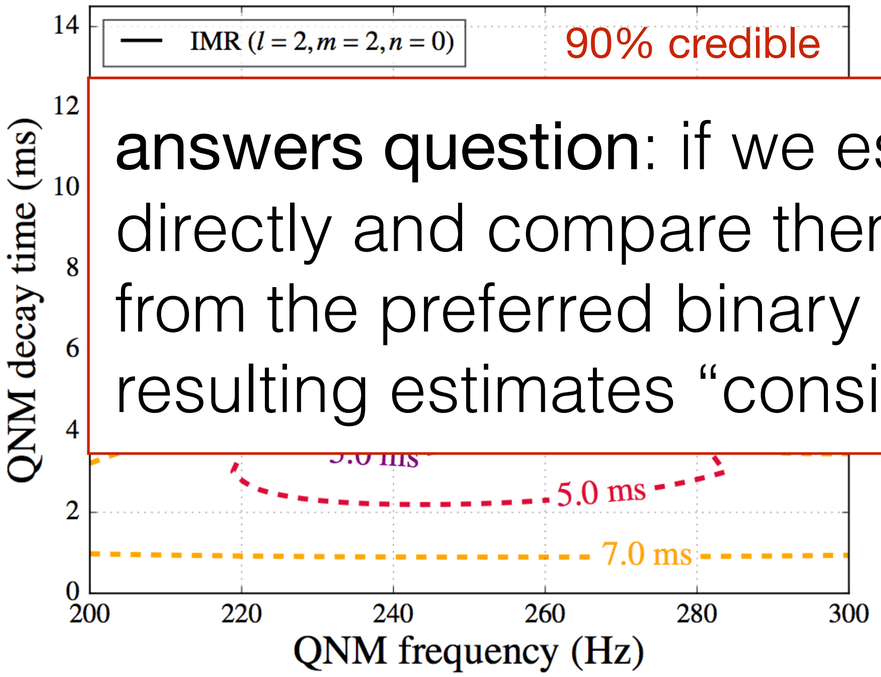
GW150914 tests of “consistency”: partial waveforms (2)



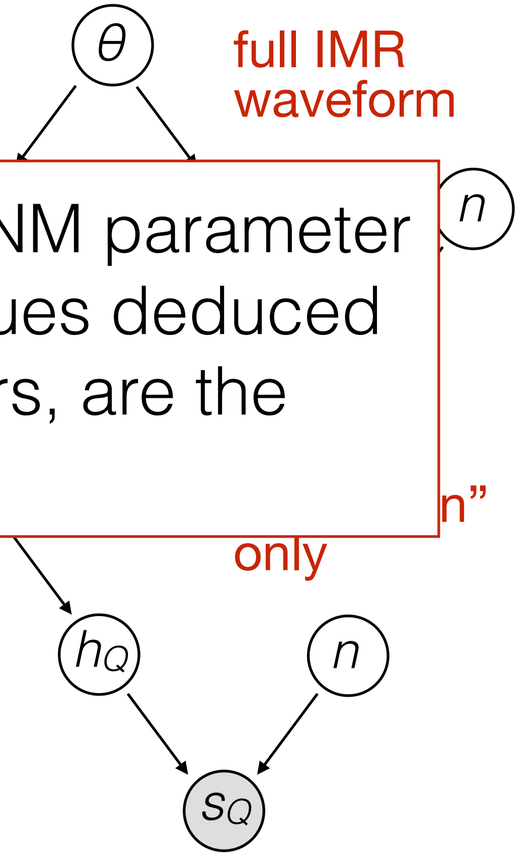


GW150914 tests of “consistency”: partial waveforms (2)

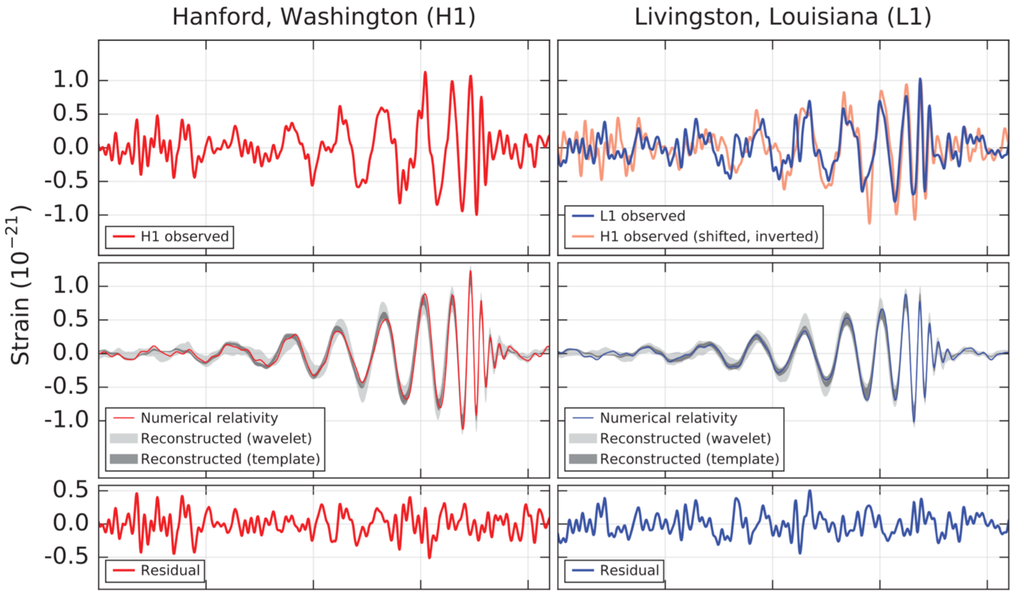
single quasi-normal mode

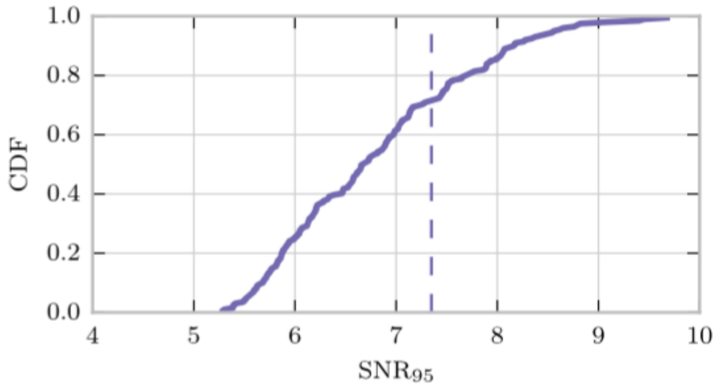


answers question: if we estimate QNM parameter directly and compare them with values deduced from the preferred binary parameters, are the resulting estimates “consistent”?



GW150914 test of “consistency”: full-waveform residual





SNR in coherent burst analysis of data residual after subtracting best-fit GW150914 waveform

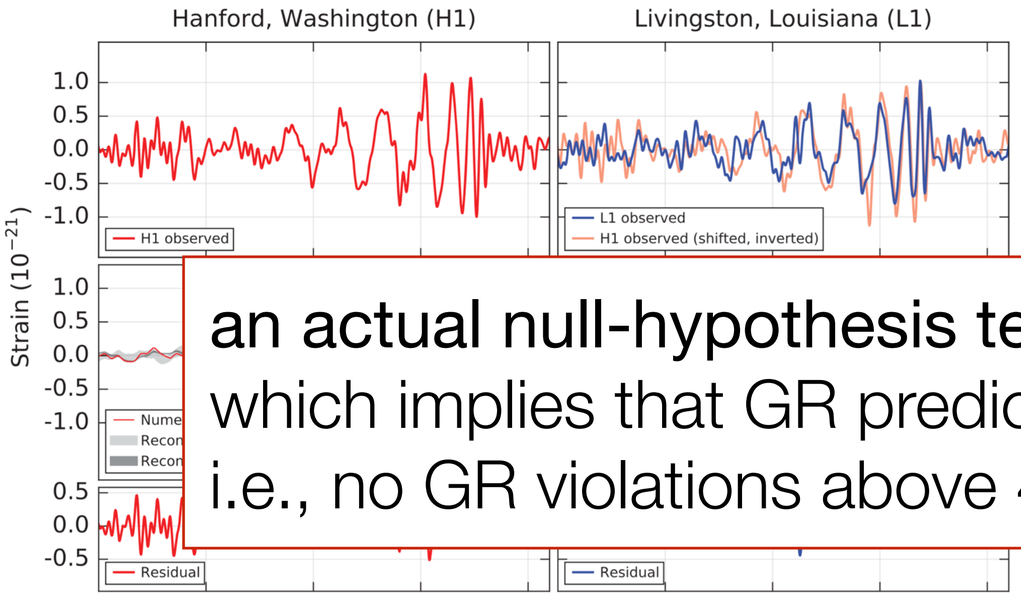
$$\mathrm{SNR}^2_{\mathrm{res}} = \frac{1-\mathrm{FF}^2}{\mathrm{FF}^2}\mathrm{SNR}^2_{\mathrm{det}}$$

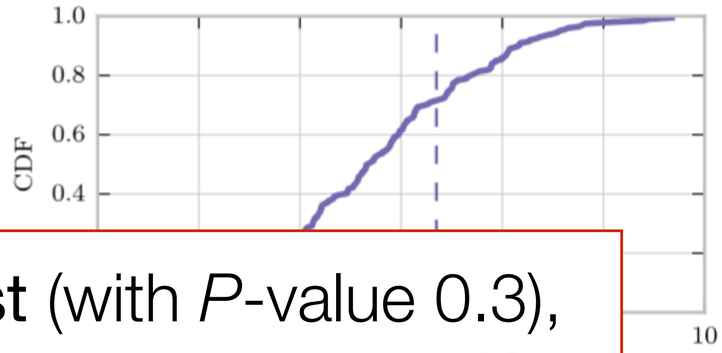
Fitting Factor: parameter-maximized waveform overlap

$$\mathrm{SNR}_{\mathrm{res}} \leq 7.3 \Rightarrow \mathrm{FF} \geq 0.96$$

(for violations not absorbed by physical parameters)

GW150914 test of “consistency”: full-waveform residual





an actual null-hypothesis test (with *P*-value 0.3), which implies that GR prediction is verified to 4%; i.e., no GR violations above 4% of waveform

best-fit GW150914 waveform

$$\mathrm{SNR}^2_{\mathrm{res}} = \frac{1-\mathrm{FF}^2}{\mathrm{FF}^2}\mathrm{SNR}^2_{\mathrm{det}}$$

Fitting Factor: parameter-maximized waveform overlap

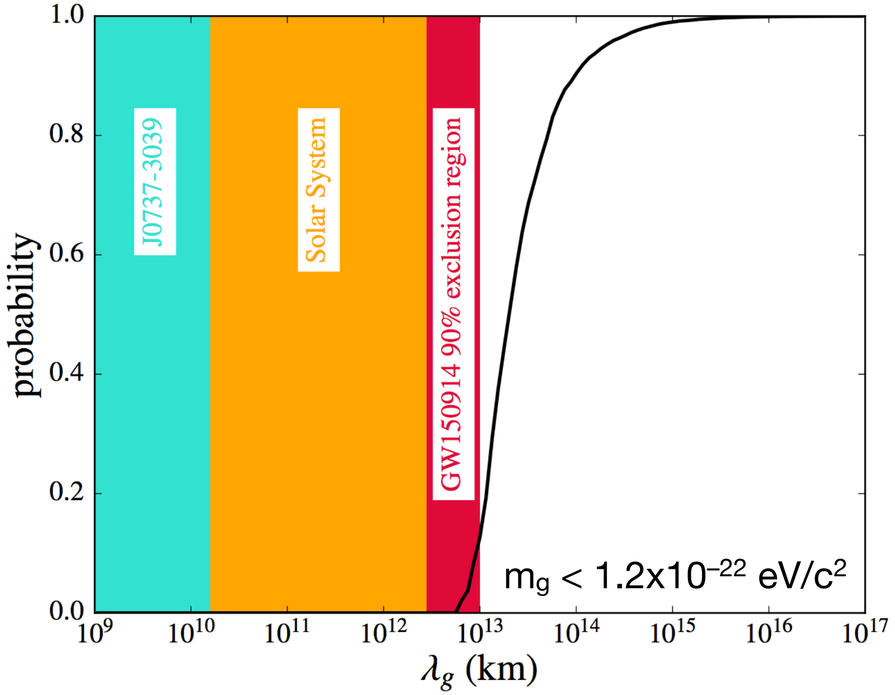
$$\mathrm{SNR}_{\mathrm{res}} \leq 7.3 \Rightarrow \mathrm{FF} \geq 0.96$$

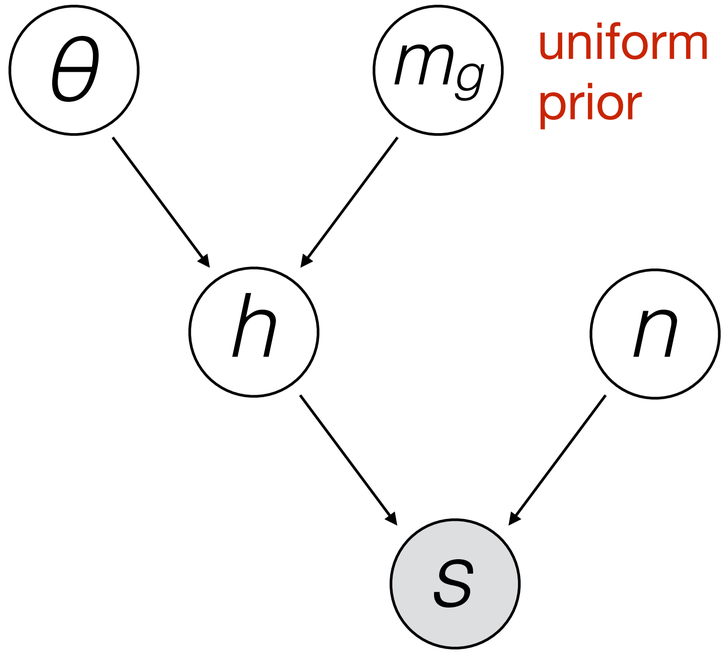
(for violations not absorbed by physical parameters)

Testing GR with GW observations: a manifesto

- **Consistency**: useful sanity checks, but hard to interpret statistically. P values are possible with much work. But would we ever believe an inconsistent result?
- **Parametric tests**: constraints on GR “constants” (PN coefficients, graviton mass) are useful proxies for increasing resolving power, but again hard to interpret. Apparent violations may focus our search for new physics.

GW150914 parametric test of GW propagation: graviton mass





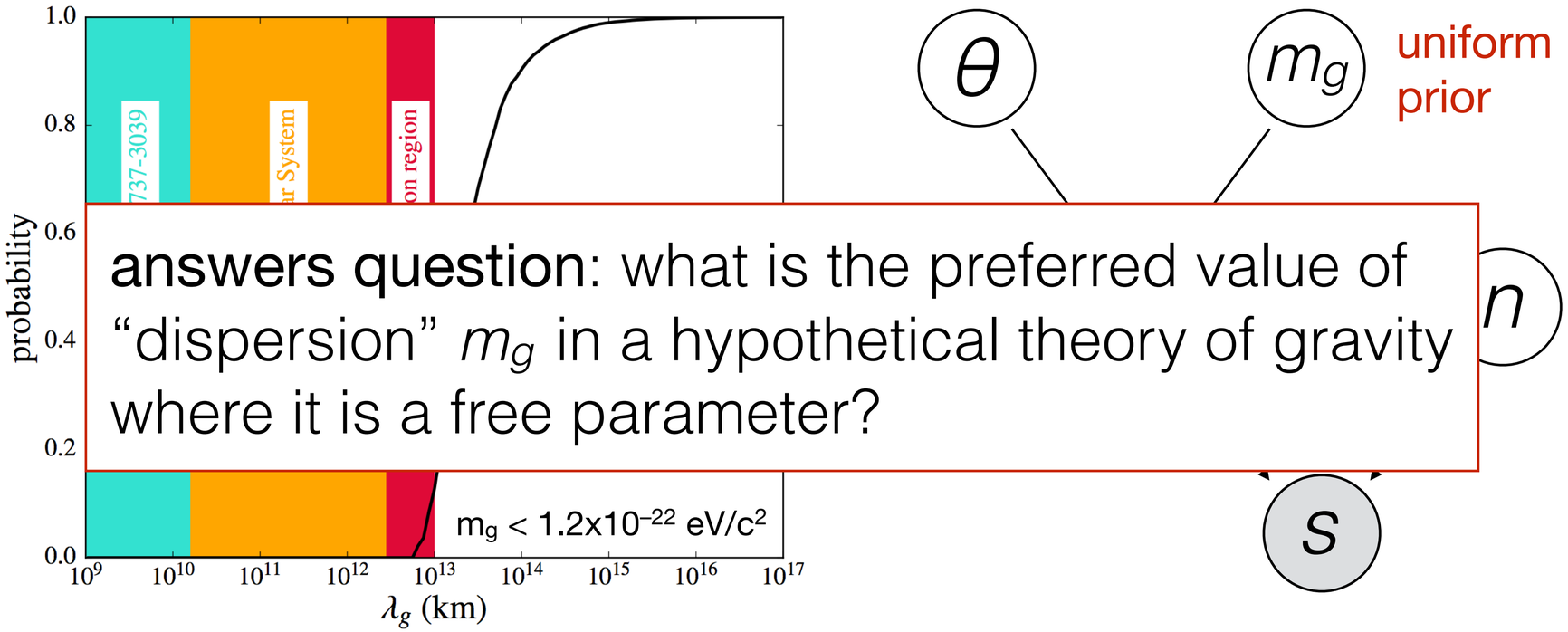
$$h(f) = \frac{1}{D}\frac{\mathcal{A}}{\sqrt{\dot{F}}} f^{2/3} e^{i\Psi(f)}$$

$$\Psi(f) = \sum_i \left[\psi_i + \psi_{il}\log f\right] f^{(i-5)/3} + \Phi^{\mathrm{MR}}[\beta_i, \alpha_i]$$

$$\frac{v_{\mathrm{g}}^2}{c^2} = 1 - \frac{m_{\mathrm{g}}^2 c^4}{E^2}$$

$$\delta\Psi(f) = \frac{\pi D c}{\lambda_g^2 (1+z) f}$$

GW150914 parametric test of GW propagation: graviton mass



answers question: what is the preferred value of "dispersion" m_g in a hypothetical theory of gravity where it is a free parameter?

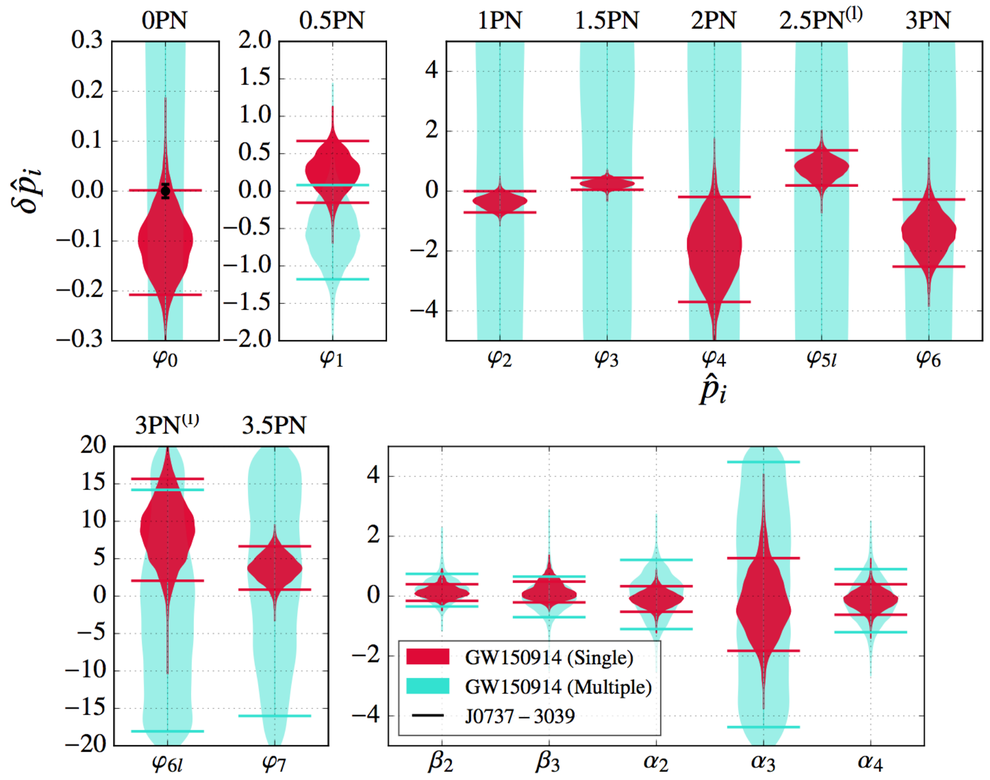
$$h(f) = \frac{1}{D}\frac{\mathcal{A}}{\sqrt{\dot{F}}} f^{2/3} e^{i\Psi(f)}$$

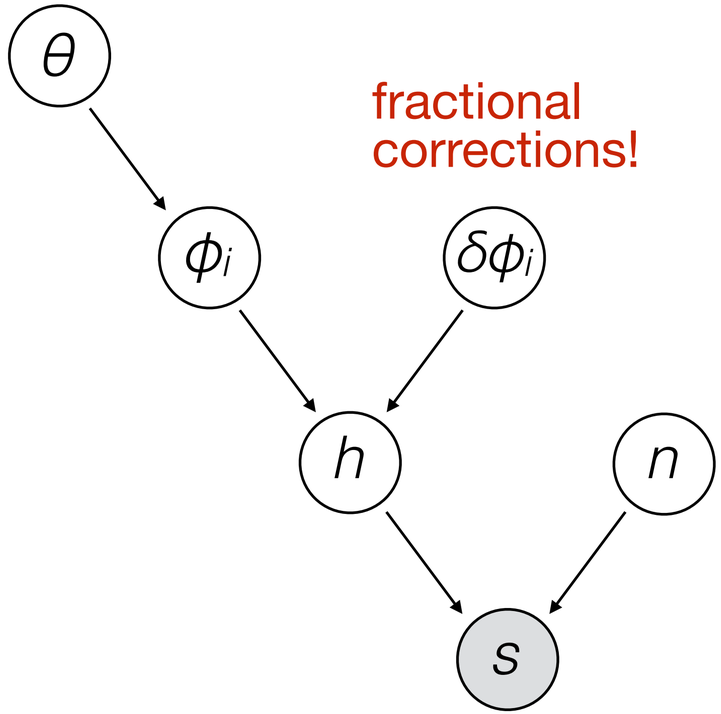
$$\Psi(f) = \sum_i \left[\psi_i + \psi_{il}\log f\right] f^{(i-5)/3} + \Phi^{\mathrm{MR}}[\beta_i, \alpha_i]$$

$$\frac{v_{\mathrm{g}}^2}{c^2} = 1 - \frac{m_{\mathrm{g}}^2 c^4}{E^2}$$

$$\delta\Psi(f) = \frac{\pi D c}{\lambda_g^2 (1+z) f}$$

GW150914 test of generation and binary dynamics: frequency-domain PN coefficients

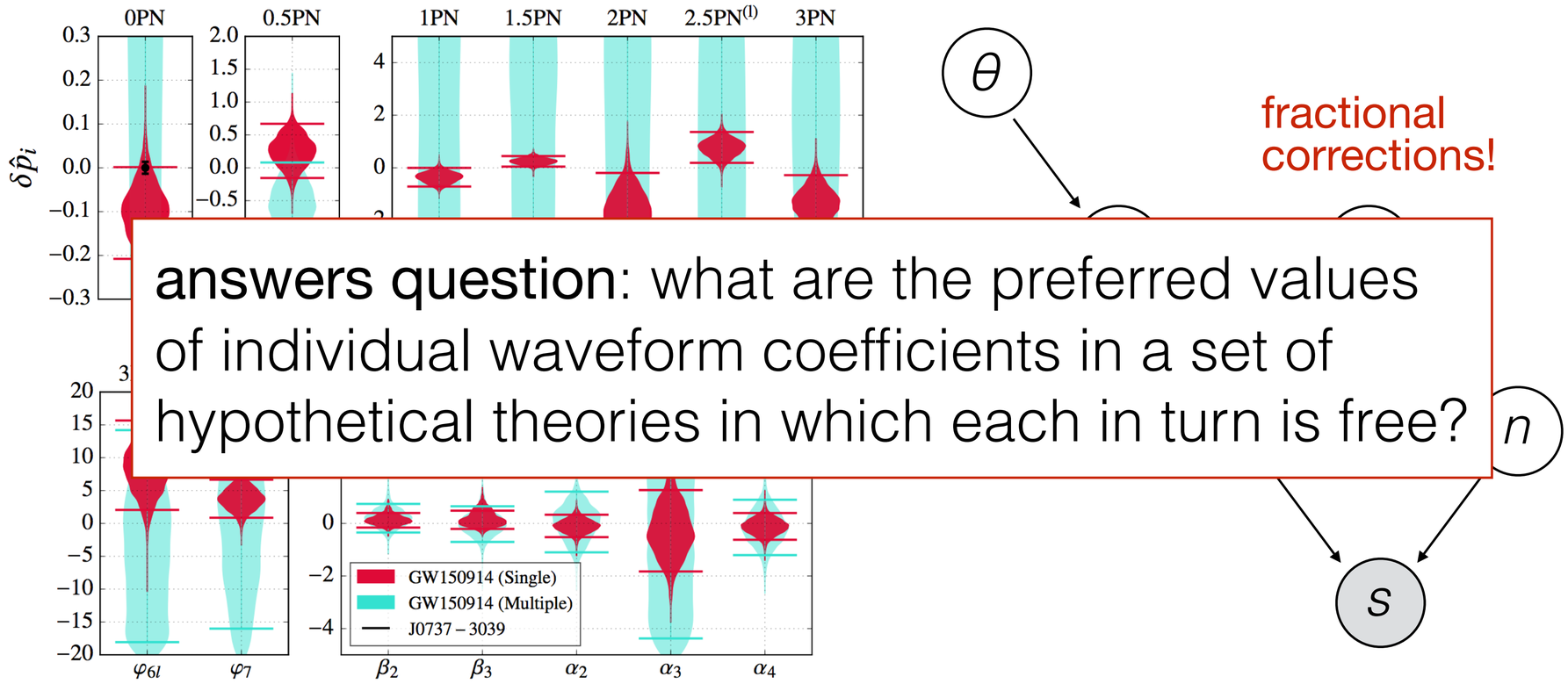




$$h(f) = \frac{1}{D}\frac{\mathcal{A}}{\sqrt{\dot{F}}} f^{2/3} e^{i\Psi(f)}$$

$$\Psi(f) = \sum_i \left[\psi_i + \psi_{il}\log f\right] f^{(i-5)/3} + \Phi^{\mathrm{MR}}[\beta_i, \alpha_i]$$

GW150914 test of generation and binary dynamics: frequency-domain PN coefficients



fractional corrections!

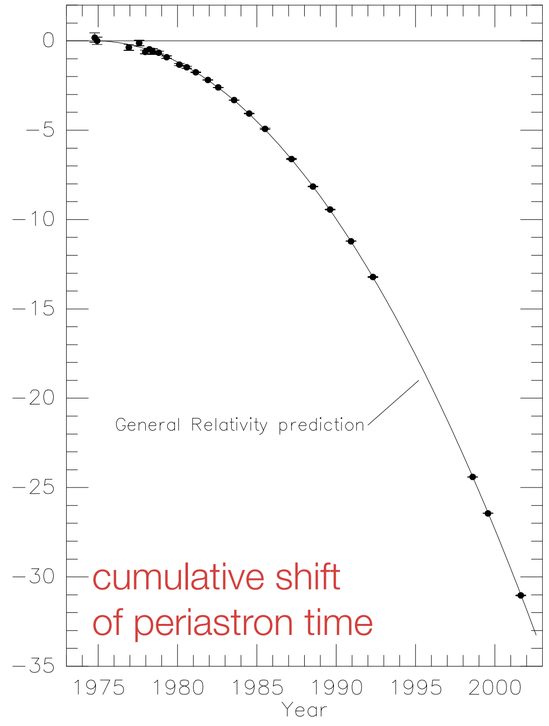
answers question: what are the preferred values of individual waveform coefficients in a set of hypothetical theories in which each in turn is free?

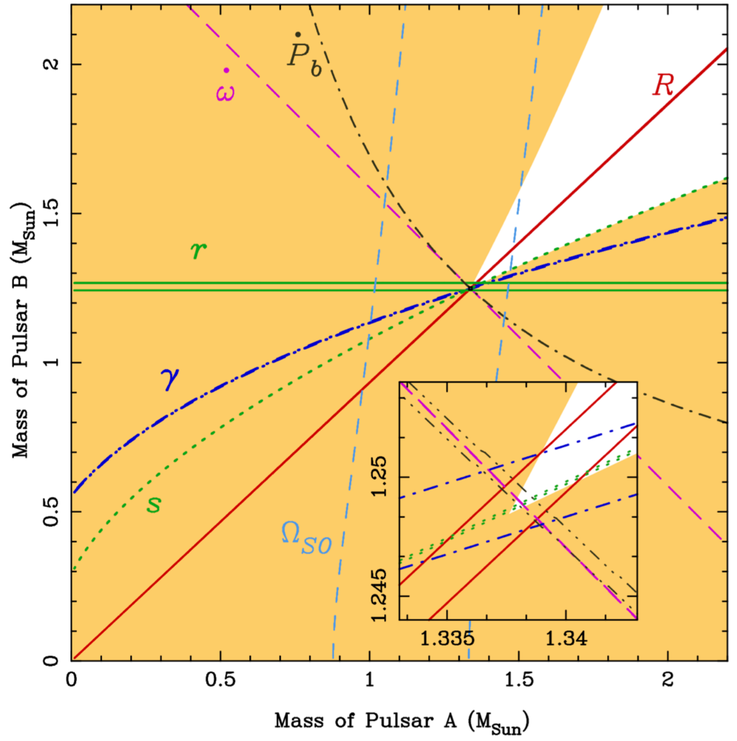
$$h(f) = \frac{1}{D}\frac{\mathcal{A}}{\sqrt{\dot{F}}} f^{2/3} e^{i\Psi(f)}$$

$$\Psi(f) = \sum_i \left[\psi_i + \psi_{il}\log f\right] f^{(i-5)/3} + \Phi^{\mathrm{MR}}[\beta_i, \alpha_i]$$

For comparison: pulsar-timing tests of GR concern physical parameters, but also have weak interpretations

PSR B1913+16
[Weisberg 2003]





PSR J0737-3039
[Kramer 2016]

$$\dot{P}_{\rm b} = -\frac{192\pi}{5}\left(\frac{P_{\rm b}}{2\pi}\right)^{-5/3}\left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right)(1-e^2)^{-7/2}\,T_\odot^{5/3}\,m_1 m_2\,M^{-1/3}$$

$$\dot{\omega} = 3\left(\frac{P_{\rm b}}{2\pi}\right)^{-5/3}(T_\odot M)^{2/3}\,(1-e^2)^{-1}, \qquad r = T_\odot\, m_2,$$

$$\gamma = e\left(\frac{P_{\rm b}}{2\pi}\right)^{1/3} T_\odot^{2/3} M^{-4/3}\, m_2(m_1+2m_2), \quad s = x\left(\frac{P_{\rm b}}{2\pi}\right)^{-2/3} T_\odot^{-1/3}\, M^{2/3}\, m_2^{-1}.$$

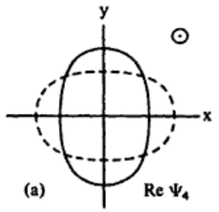
Testing GR with GW observations: a manifesto

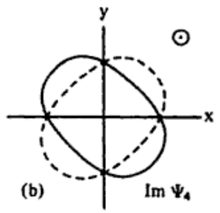
- **Consistency**: useful sanity checks, but hard to interpret statistically. P values are possible with much work. But would we ever believe an inconsistent result?
- **Parametric tests**: constraints on GR “constants” (PN coefficients, graviton mass) are useful proxies for increasing resolving power, but again hard to interpret. Apparent violations may focus our search for new physics.
- **Alternative theories**: new physics will be established by model comparison of GR with fully predictive alternative theories. It is a problem to establish Bayesian priors for alternative gravity, and for alternative-gravity parameters.

same as detection

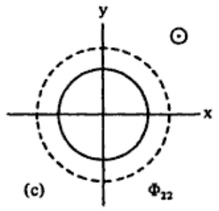
other measurements, particle physics, astroparticles

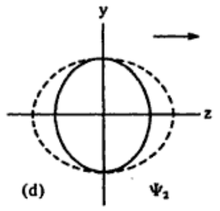
GW150914 test of polarization



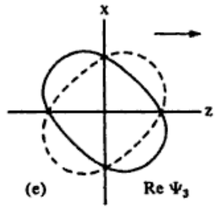


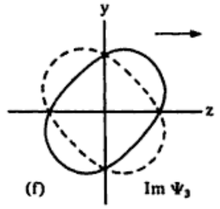
Tensor



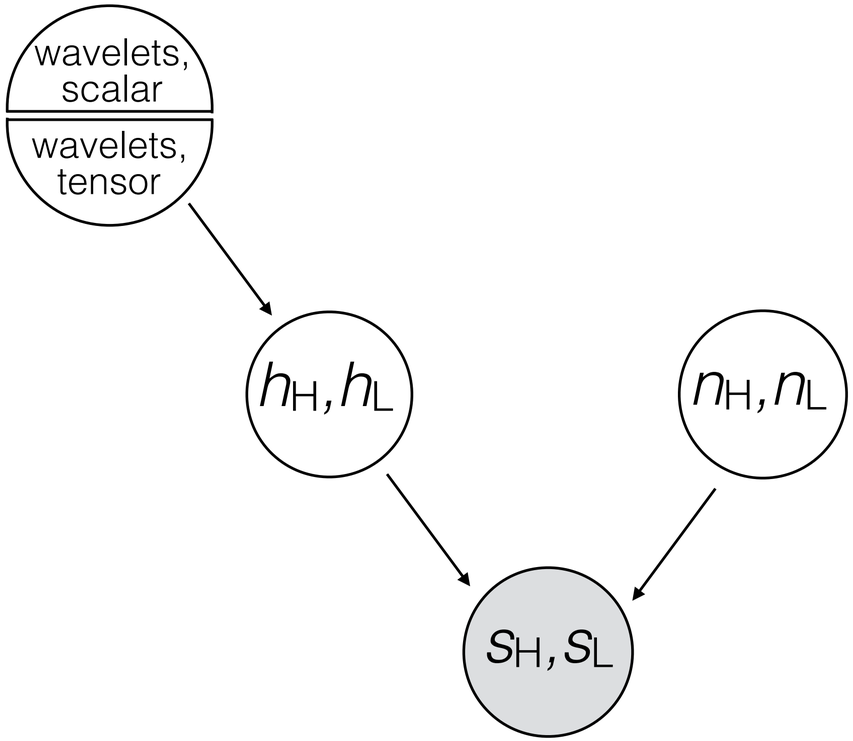


Scalar





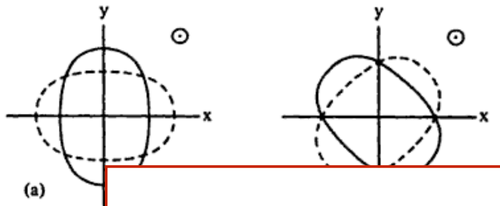
Vector



no discrimination between S and T

$$\frac{P(s|\mathrm{S})}{P(s|\mathrm{T})} \sim 1$$

GW150914 test of polarization

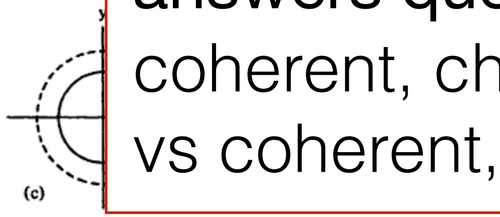


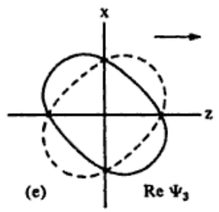
Tensor

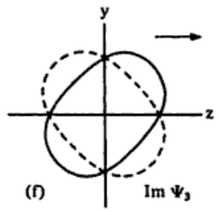
wavelets, scalar

wavelets, tensor

answers question: is there evidence for coherent, chirp-like quadrupolar radiation vs coherent, chirp-like purely scalar radiation?







Vector

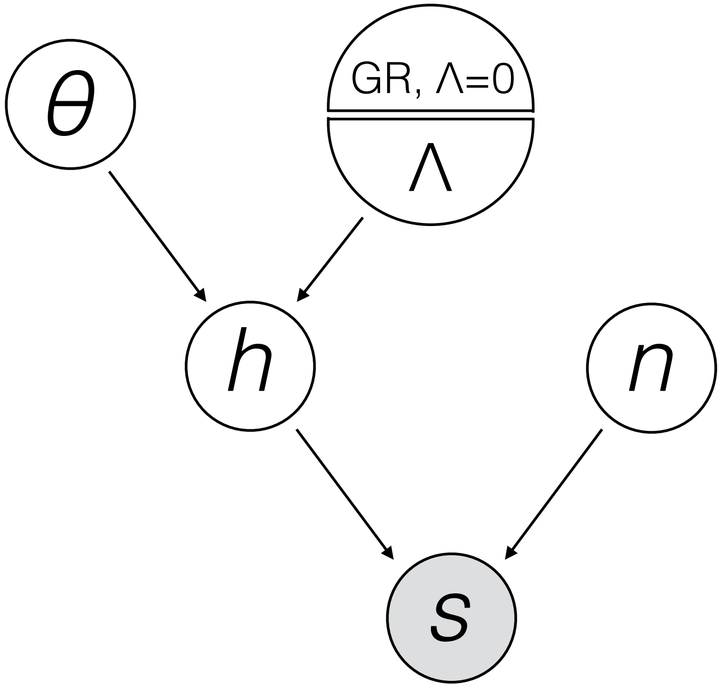
n_H, n_L

s_H, s_L

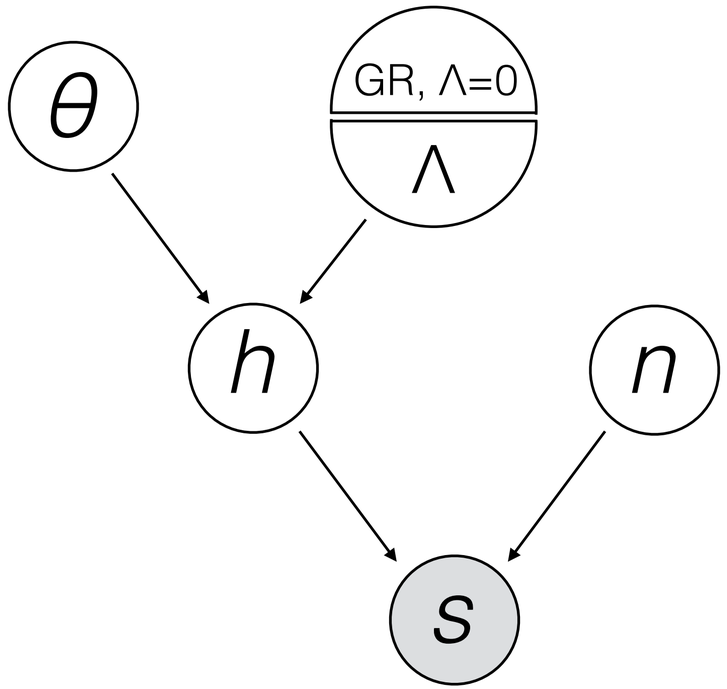
no discrimination between S and T

$$\frac{P(s|\mathrm{S})}{P(s|\mathrm{T})} \sim 1$$

Establishing alternative theories



Establishing alternative theories



new physics follows from establishing an **anomaly**: we need to obtain convincing evidence that the data prefers an alternative theory of gravity over GR

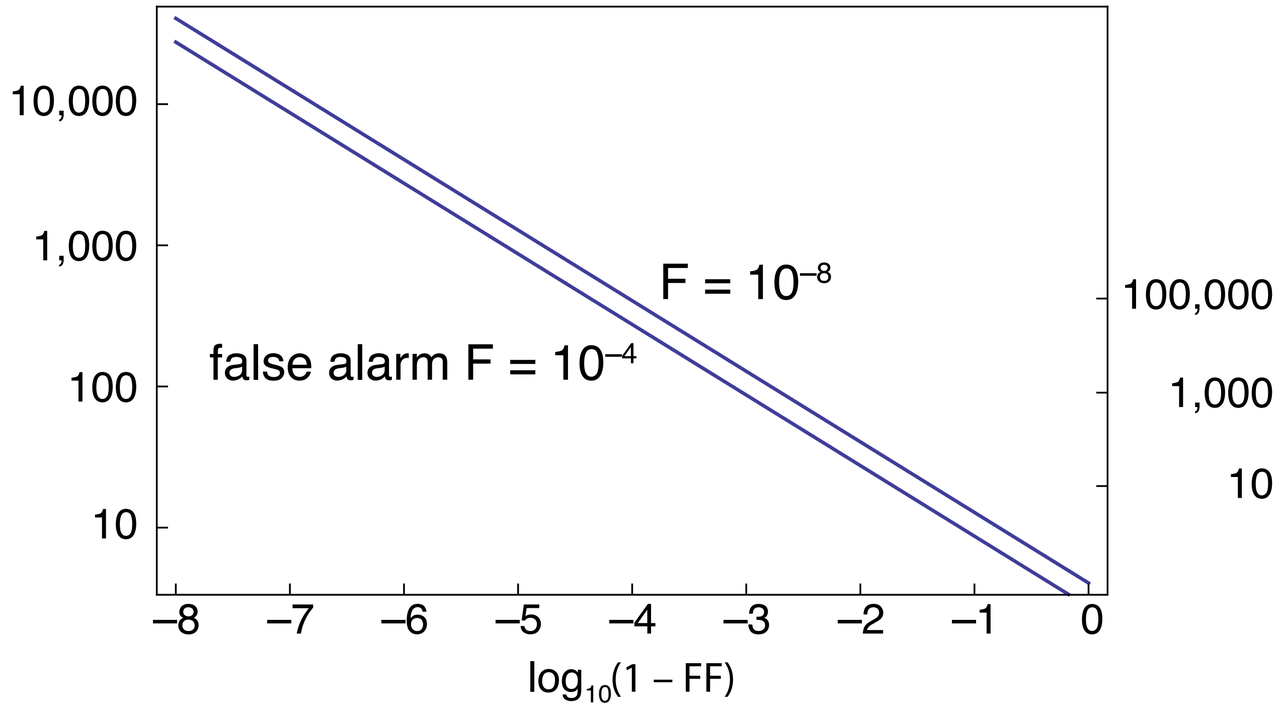
Testing GR with GW observations: a manifesto

- **Consistency**: useful sanity checks, hard to interpret statistically. P values are possible with much work. But would we ever believe an inconsistent result?
- **Parametric tests**: constraints on GR “constants” (PN coefficients, graviton mass) are useful proxies for increasing resolving power, but again hard to interpret. Apparent violations may focus our search for new physics.
- **Alternative theories**: new physics will be established by model comparison of GR with fully predictive alternative theories. It is a problem to establish Bayesian priors for alternative gravity, and for alternative-gravity parameters.
- **Size of effects**: detection SNR determines the magnitude of detectable waveform anomalies. 1% LIGO, up to 10^{-5} for LISA and future ground-based detectors.

How well can we hope to do? [MV, PRD 86, 2012]

- For a fixed false-alarm rate, we ask what **SNR** is needed to detect AG with 50% probability as a function of **fitting factor FF**, using the Bayesian odds ratio as “detection” statistic.

SNR to detect AG



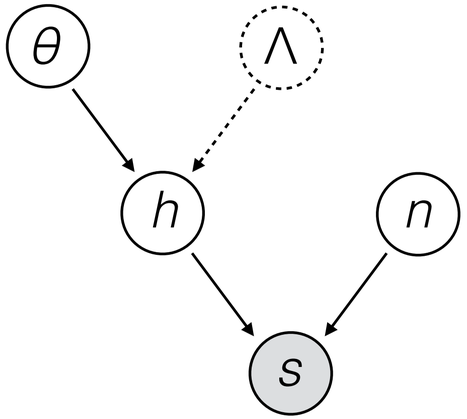
events to yield SNR

Testing GR with GW observations: a manifesto

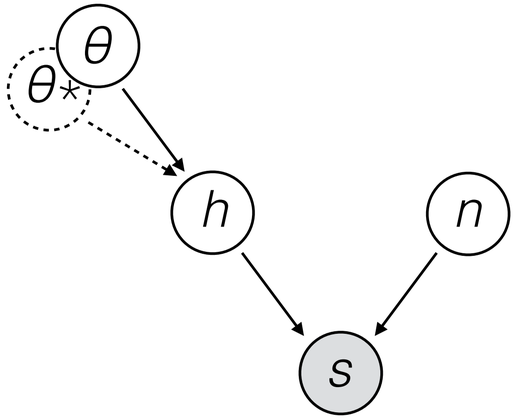
- **Consistency**: useful sanity checks, hard to interpret statistically. P values are possible with much work. But would we ever believe an inconsistent result?
- **Parametric tests**: constraints on GR “constants” (PN coefficients, graviton mass) are useful proxies for increasing resolving power, but again hard to interpret. Apparent violations may focus our search for new physics.
- **Alternative theories**: new physics will be established by model comparison of GR with fully predictive alternative theories. It is a problem to establish Bayesian priors for alternative gravity, and for alternative-gravity parameters.
- **Size of effects**: detection SNR determines the magnitude of detectable waveform anomalies: 1% LIGO, up to 10^{-5} for LISA and future ground-based detectors.
- **Systematics**: beyond statistical-significance arguments, we will need a solid chain of evidence before we claim GR is violated.

Control of systematics

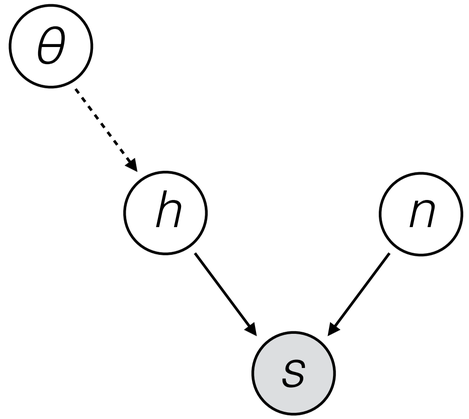
fundamental bias (Yunes)



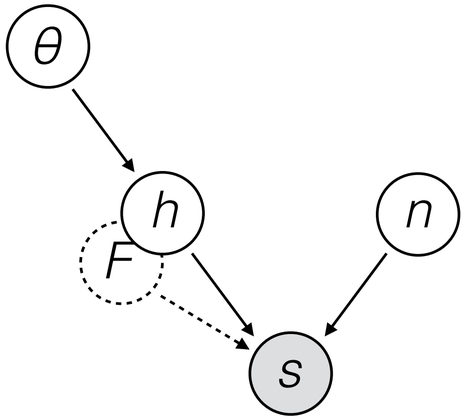
astrophysical bias



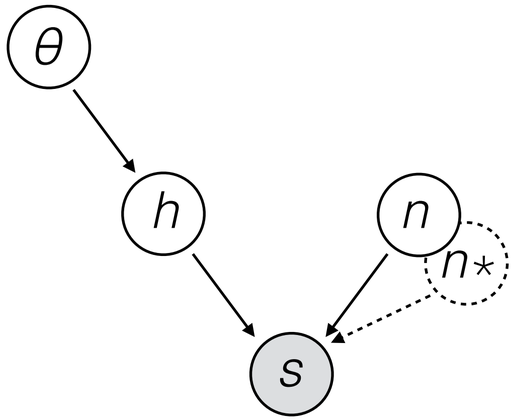
mismodeling



calibration



glitches



Testing GR with GW observations: recommendations

- **Consistency**: useful sanity checks, hard to interpret statistically. P values are possible with much work. But would we ever believe an inconsistent result?
- **Parametric tests**: constraints on GR “constants” (PN coefficients, graviton mass) are useful proxies for increasing resolving power, but again hard to interpret. Apparent violations may focus our search for new physics.
- **Alternative theories**: new physics will be established by model comparison of GR with fully predictive alternative theories. It is a problem to establish Bayesian priors for alternative gravity, and for alternative-gravity parameters.

develop!

- **Size of effects**: detection SNR determines the magnitude of detectable waveform anomalies: 1% LIGO, up to 10^{-5} for LISA and future ground-based detectors.
- **Systematics**: beyond statistical-significance arguments, we will need a solid chain of evidence before we claim GR is violated.

Testing GR with GW observations: recommendations

- **Consistency**: useful sanity checks, hard to interpret statistically. P values are possible with much work. But would we ever believe an inconsistent result?
- **Parametric tests**: constraints on GR “constants” (PN coefficients, graviton mass) are useful proxies for increasing resolving power, but again hard to interpret. Apparent violations may focus our search for new physics.
- develop! **Alternative theories**: new physics will be established by model comparison of GR with fully predictive alternative theories. It is a problem to establish Bayesian priors for alternative gravity, and for alternative-gravity parameters.
- fund! **Size of effects**: detection SNR determines the magnitude of detectable waveform anomalies: 1% LIGO, up to 10^{-5} for LISA and future ground-based detectors.
- **Systematics**: beyond statistical-significance arguments, we will need a solid chain of evidence before we claim GR is violated.

Testing GR with GW observations: recommendations

- **Consistency**: useful sanity checks, hard to interpret statistically. P values are possible with much work. But would we ever believe an inconsistent result?
- **Parametric tests**: constraints on GR “constants” (PN coefficients, graviton mass) are useful proxies for increasing resolving power, but again hard to interpret. Apparent violations may focus our search for new physics.

develop!

- **Alternative theories**: new physics will be established by model comparison of GR with fully predictive alternative theories. It is a problem to establish Bayesian priors for alternative gravity, and for alternative-gravity parameters.

fund!

- **Size of effects**: detection SNR determines the magnitude of detectable waveform anomalies: 1% LIGO, up to 10^{-5} for LISA and future ground-based detectors.

control!

- **Systematics**: beyond statistical-significance arguments, we will need a solid chain of evidence before we claim GR is violated.

Testing GR with GW observations: recommendations

- **Consistency**: useful sanity checks, hard to interpret statistically. P values are possible with much work. But would we ever believe an inconsistent result?
- **Parametric tests**: constraints on GR "constants" (PN coefficients, graviton mass) are useful proxies for increasing resolving power, but again hard to interpret. Apparent violations may focus our search for new physics.
- **Alternative theories**: new physics will be established by model comparison of GR with fully predictive alternative theories. It is a problem to establish Bayesian priors for alternative gravity, and for alternative-gravity parameters.
- **Size of effects**: detection SNR determines the magnitude of detectable waveform anomalies: 1% LIGO, up to 10^{-5} for LISA and future ground-based detectors.
- **Systematics**: beyond statistical-significance arguments, we will need a solid chain of evidence before we claim GR is violated.

fold!

develop!

fund!

control!